

## On the hydrodynamics of swimming enzymes

Xiaoyu Bai<sup>1, a)</sup> and Peter G. Wolynes<sup>1, b)</sup>

*Department of Chemistry and Center for Theoretical Biological Physics,  
Rice university, Houston TX 77005*

(Dated: 8 October 2015)

Several recent experiments suggest that rather generally the diffusion of enzymes may be augmented through their activity. We demonstrate that such swimming motility can emerge from the interplay between the enzyme energy landscape and the hydrodynamic coupling of the enzyme to its environment. Swimming thus occurs during the transit time of a transient allosteric change. We estimate the velocity during the transition. The analysis of such a swimming motion suggests the final stroke size is limited by the hydrodynamic size of the enzyme. This limit is quite a bit smaller than the values that can be inferred from the recent experiments. We also show that one proposed explanation of the experiments based on reaction heat effects can be ruled out using an extended hydrodynamic analysis. These results lead us to propose an alternate explanation of the fluorescence correlation measurements.

---

<sup>a)</sup>Xiaoyu.Bai@rice.edu

<sup>b)</sup>pwolynes@rice.edu

# I. INTRODUCTION

All swimming ultimately can be traced to the dynamics of enzymes. Both animal muscles and the cytoskeletons of individual cells move through the cooperation of many motor proteins, which are enzymes that act together in large scale structures<sup>1,2</sup>. Can individual enzymes, however, swim? How would they do so? The possibility of single particle molecular locomotion was already contemplated in setting up the theory of motorized crystals, motorized glasses and active molecular matter years ago<sup>3-9</sup>. Nevertheless, we were surprised by recent observations<sup>10,11</sup> that suggest that a large range of enzymes, most of which are not in any way involved in biological motor activity, appear to swim, albeit in an undirected manner. Two groups have reported enhanced diffusion of several different enzymes that include catalase, urease and alkaline phosphatase. None of these are classical motor proteins. The apparent enhancement of diffusion appears to be proportional to enzymatic activity, just as is predicted by the theory of motorized assemblies<sup>3,5</sup>. In their largely observational paper Riedel et al.<sup>10</sup> also suggested a schematic mechanism by which the internal chemical energy in the substrate-enzyme complex could be transduced into motion of the enzyme's center of mass. They postulated the idea that the heat released by the reaction would lead to a pressure impulse in the surrounding water that in turn would lead to the motion of enzyme as a whole.

The reaction heat hypothesis was apparently inspired by the experimental observation that they found no enhanced diffusion for the enzyme triose phosphate isomerase, which catalyzes a reaction that does not release heat, while the other enzymes that catalyze reactions with larger  $\Delta H$ 's did apparently display activity enhanced diffusion. Although the quantitative details of the proposed locomotion mechanism were not completely laid out, on its face, the heat hypothesis itself raises some questions. For instance, if even momentarily the motion of the protein center of mass is supposed to be directed, why should the enthalpy change be the relevant thermodynamic quantity for determining the impulse, rather than a free-energy change? Also well established arguments suggest that the large scale motions of proteins should be highly damped by the solvent<sup>12,13</sup> and that therefore such motions of proteins should be described by the hydrodynamics of bodies at low Reynolds number and with low Mach number, the ratio of the characteristic speed of the moving object and the

sound speed in the medium. The arguments put forward by Riedel et al. rely on the finite compressibility of the surrounding solvent and thus their picture entails high Mach number hydrodynamics.

In this paper we first explore an alternative explanation for how single enzymes might be able to swim while nevertheless only moving at low Reynolds and Mach numbers. The general problem of swimming at such low speeds has formed a long standing elegant part of biologically inspired physics starting with Purcell’s so-called ”Scallop Theorem”<sup>14</sup>. He used the theorem to point out the difficulties bacteria face in swimming. This theorem was later rigorously proved by Shapere and Wilczek<sup>15</sup>. The Scallop theorem states that to swim at low Reynolds number, the swim cycle must involve changing at least two degrees of freedom and also that the sequence of changes must not be time reversal invariant. The Scallop theorem constraints arise because the incompressible steady Navier-Stokes equations that describe fluid motion at low Reynolds number are time reversal invariant: reversing a forward stroke therefore causes the swimmer simply to re-trace its forward motion in the reverse sense so as to yield no net displacement<sup>16,17</sup>. Nevertheless, a cyclic motion in two or more degrees of freedom need not be time reversal invariant so that asymmetric cyclic movements can allow a swimmer to crawl through its surroundings, no matter how viscous they are. Since enzymes are complicated molecules with many more than two degrees of freedom, and also recognizing that the enzymes degrees of freedom can be restored to equilibrium after release of a catalytic product without actually reversing the catalysis step itself, the Scallop Theorem does indeed allow an enzyme to translate forward in the process of carrying out a series of chemical reactions. Treating this problem for any specific enzyme in structural detail would doubtless be quite complex so here to make our conceptual point we confine ourselves to studying a very schematic model that envisions directly coupling the reaction coordinate of the reaction itself or a subsequent enzyme allosteric structural change coordinate describing motion after an enzymatic reaction to other degrees of freedom of the enzyme molecule that describe the relative motion of domains in the protein. While accounting explicitly for the dynamic stochastic coupling between reaction modes and the overall enzyme motion, this model is quite parallel to models of nanomotors already put forward that use deterministic cycles of swimmer shape change<sup>18</sup>. We see in this model that a key role is played by the hydrodynamics of coupling during the traversal period of

the activated motion. The latter quantity has recently received much attention through experiments on protein folding<sup>26,27</sup>. Even the simplest form of the model, in our view, gives a plausible upper limit to the possible speed and stroke size of an enzyme moving through a solution no matter what the structural details of the enzyme cycle. In the most favorable imaginable case the maximum predicted stroke size from this estimate turns out to be smaller than the enzyme's hydrodynamic radius. We can easily imagine by involving more elaborate motional mechanisms that the final stroke size could actually be much smaller than the enzyme's hydrodynamic size, but as we see it, achieving stroke sizes significantly larger than the hydrodynamic radius would require a huge conformation change of the enzyme tantamount to its global folding. While actual biological motor motions are known to involve "cracking"<sup>19</sup>, nothing as dramatic as complete unfolding has yet been contemplated, even for motor proteins, and indeed such unfolding seems even more unlikely for the particular enzymes that were studied experimentally, which are quite stable. The upper bound character of our result therefore raises some difficulties for the present interpretation of the experiments: while a step size as big as the hydrodynamic radius is quite substantial and would clearly be adequate for most biological functional purposes, our predicted bound is much too small to explain the reported enhancements of diffusion; indeed even the maximum step size that our model would predict gives enhancements decidedly below the observability limits for the fluorescence techniques employed. The reported values of the diffusion enhancement for catalase in particular requires a step roughly 6 times larger than catalase's hydrodynamic radius. Clearly our theory cannot account for a step size with the large value inferred from the experiments. As we shall show explicitly it is also hard to see how high Mach numbers can ever be achieved during enzyme locomotion so we also cannot see how the explanation based on the heating mechanism can be hydrodynamically plausible. Rather different mechanisms for enzyme locomotion have been proposed such as self-electrophoresis<sup>20</sup>. These mechanisms involve transiently modifying the composition of the solvent. Such effects are available for larger objects<sup>16-18</sup> but these models, however, have subsequently been withdrawn as explanations for enzyme locomotion by their authors. At the end of this paper, we are therefore led to suggest an alternative interpretation of the fluorescence experiments whereby the measured changes in fluorescence do not in fact arise from enhanced diffusion due to enzyme swimming at all but rather arise from chemically specific sources of transient fluorescence quenching that have not been taken into account.

## II. THE SWIMMING ENZYME MODEL

Our schematic model of a swimming enzyme envisions three protein domains that are hydrodynamically coupled to their surroundings. The relative motions of the domains are governed by two different free energy surfaces. One surface that describes the conformational changes accompanying ligand hydrolysis is called the "excited state" surface  $F_e$  while the other, the so-called ground state surface  $F_g$ , describes the motion accompanying ligand binding and enzyme structural relaxation in the absence of hydrolysis. When these free energy surfaces are projected on the domain locations  $r_1$ ,  $r_2$  and  $r_3$  we take the excited surface  $F_e$  to be a bistable function of one of the interdomain distances, say  $x_1$ , as shown in Fig. 1 while the relaxational ground state surface  $F_g$  is taken to have a single minimum and is effectively harmonic. For simplicity we do not treat explicitly motion on the ground state surface and thus suppose the locomotion primarily occurs during the hydrolysis or allosteric step in the upper excited surface. Other possible schemes can be treated in a similar fashion to the present analysis. This set up is motivated by conformation switching models introduced earlier to describe allostery<sup>21</sup>. As pictured in Fig. 1 the actual hydrolysis and release events are supposed to occur instantaneously so that the domains are considered to be fixed during the bond breaking events but the relevant motions of the domains after the chemical changes are generated by motion on either of the two surfaces  $F_g(x_1, x_2)$  and  $F_e(x_1, x_2)$ . The dynamics of the domain coordinates are coupled to the solvent hydrodynamically following the original scheme of Najafi and Golestanian<sup>18</sup>. Again for simplicity we take the domains to be equal size spheres with radius  $a$  that move one dimensionally along their mutual axis. The full cycle of motion therefore consists of stochastic instantaneous switching between these two surfaces interspersed with hydrodynamically coupled Brownian motion upon these two distinct surfaces. This Brownian motion involves the internal motion of the allosteric protein that ends up being frictionally coupled by hydrodynamics to the overall displacement of the enzyme.

The actual swimming motion occurs on the excited state surface by the internal coordinate of the enzyme stochastically leaping up from the initial configuration to the transition state configuration and then falling down to a product configuration. While making this leap fluid is displaced and the center of mass of the enzyme thereby moves. The most probable

transition path can be found using by a path integral treatment that amounts to calculating the average of a path ensemble made up of all possible Brownian traversals over the free energy barrier. The details of this argument may be found in the Appendix. If we picture the bistable potential as being an inverted harmonic well with a spring constraint of  $-k(k > 0)$  the motions involved in traversing the free energy barrier can be found explicitly. Assuming equal and opposite negative spring constants for each domain as well as equilibrium spring length  $x$  for each domain the excited state barrier potential can be written as  $F_e(x_1, x_2) = -\frac{1}{2}k(x_1 - x)^2 + \frac{1}{2}k(x_2 - x)^2$ . The hydrodynamically coupled Brownian Dynamics equations for the model swimmer in matrix form are:

$$\frac{dr_i}{dt} = \beta \sum_j D_{ij} F_j + \eta_i(t) \quad (1)$$

where  $F_i$  is the effective spring force exerted on each sphere and is determined by the excited potential  $F_e$ .  $\eta_i(t)$  are white noise terms that are related to diffusion tensor  $D_{ij}$  by the fluctuation-dissipation theorem  $\langle \eta_i(t) \eta_j(t') \rangle = 2D_{ij} \delta(t - t')$ . Since the hydrodynamic coupling is kept at Oseen level as the first matrix inside the square bracket on the r.h.s. in equation (1), the multiparticle diffusion tensor can be written as  $D_{ij} = k_B T [\delta_{ij} \zeta_0^{-1} + (1 - \delta_{ij}) T(r_{ij})]$ , where  $T(r) = (8\pi\eta_0 r)^{-1} (I + \frac{\hat{r}\hat{r}}{r^2})$  is the Oseen Tensor<sup>22</sup>. To solve the coupled equations we carry out an eigen-analysis of the deterministic part for these linear equations. This is equivalent to extremizing the Onsager-Machlup Lagrangian so as to allow us to obtain the most probable path for a traversal from one well to the other. Since three degrees of freedom are involved we find three eigenvalues. The eigenvalues for motion across the barrier consist of one zero eigenvalue  $\lambda_1 = 0$  and two nonzero values: one stable  $\lambda_2 = -\sqrt{3}k(1 - \frac{5a}{4x})/\zeta_0$  and the other unstable  $\lambda_3 = \sqrt{3}k(1 - \frac{5a}{4x})/\zeta_0$ . The existence of the zero frequency collective mode demonstrates the capability of net locomotion of the model swimmer. The dominant path for the collective mode (swimming motion) arises from this model and becomes simply:

$$r'_1(t) = \frac{1}{3}[r_1(t) + r_2(t) + r_3(t)] = -\frac{7ka^2}{8\zeta_0 x}t + x \quad (2)$$

where the initial positions of the three spheres are set as 0,  $x$  and  $2x$  respectively. We see the collective mode  $r'_1$  with the zero eigenvalue yields the motion of center of mass of the swimmer starting at location  $x$  as a uniform motion. Dominant paths for the auxiliary domains  $r'_2(t)$  and  $r'_3(t)$  are discussed in the SI. We also see that the swimming velocity

during the transit time is  $v_{swim} = -\frac{7ka^2}{8\zeta_0 x}$  clearly showing the hydrodynamic coupling of the swimmer to the solvent is crucial to its swimming behavior. How long does a given traversal take? This is essentially the time for the internal coordinate to go from the bottom of one well to the other. The transit time is obviously a distributed property. For the inverted oscillation model it is essentially the time for the reactive mode to move sufficiently to change the energy by an amount of  $\Delta U^\ddagger$ . The mean transit time therefore increases logarithmically with the increase of barrier height  $\Delta U^\ddagger = \frac{1}{2}m\omega^\ddagger x^\ddagger{}^2$  since the internal coordinate shift on the inverted potential grows exponentially in time being described as the motion of an unstable damped system describing the domain motion. Typical traversal times have been discussed for the simplest models theoretically<sup>23–25</sup> and indeed have been measured experimentally for biomolecular folding process<sup>26,27</sup>. While the dominant path argument gives the most probable traversal time, one can also find explicitly the distribution for the transit time by normalizing the probability density flux from the pre-stroke to the post-stroke state with the flux being determined by the transit probability. This distribution can also be explicitly obtained from our path integral treatment. In the Appendix we provide the details of the derivation (see Appendix). The mean traversal time  $T$  determined from the distribution turns out to be a bit larger than the most probable transit time. Explicitly we find:

$$T \approx \frac{\ln(2\beta\Delta U^\ddagger) + \gamma + \ln[\sqrt{3}(1 - \frac{5a}{4x})/\cos^2 \phi]}{\lambda_3} \quad (3)$$

where  $\gamma$  is the Euler constant and  $\phi$  is the angle between the reaction coordinate and the eigenvector associated with  $\lambda_3$  in the euclidean space spanned by  $r_1$ ,  $r_2$  and  $r_3$ . This angle follows from the hydrodynamic coupling as described in the Appendix. Knowledge of the free energy barrier and this eigenvalue corresponding to the unstable mode is enough for evaluating the mean transit time  $T$ , but determining the coupled dominant leaping path demands the motions in the pre-stroke and post-stroke state of enzymes, which we neglect here. A simple direct prediction for the stroke size  $L$  possible at the Oseen limit ( $\frac{a}{x} \rightarrow 0$ ) follows from this estimation for the mean transit time combined with the velocity of the collective mode that results from the path-integral treatment:

$$L = |v_{swim} \times T| = \frac{7a^2[\ln(2\beta\Delta U^\ddagger) + \gamma + \ln(\sqrt{3}/\cos^2 \phi)]}{8\sqrt{3}x} \quad (4)$$

The Euler constant  $\gamma$  is small and the free energy barrier for a typical enzymatic reaction is just several  $k_B T$ . For most enzymes, the size of each subunit  $a$  is usually comparable to

the separation distance  $x$  and thus, one can infer from the equation (4) that the size of the net displacement can only be of the same order as the hydrodynamic size of enzymes, e.g.  $L \sim \frac{a^2}{x} \sim a$ . From this formula one can estimate the maximum stroke sizes. We do this for catalase, one of the four enzymes upon which the fluorescence correlation measurement have been conducted by Riedel et al<sup>10</sup>. The hydrodynamic radius of catalase is 5.22nm. The stroke sizes for catalase assuming three different Oseen ratios  $\frac{a}{x} = 0.1, 0.5$  and 1 turn out to be correspondingly  $9.2 \times 10^{-2}$ nm, 2.3nm and 9.2nm where  $\Delta U^\ddagger$  is taken as  $5k_B T$ . Riedel et al. did not report stroke size, but the theory of motorized assemblies suggests that the slope  $\alpha$  of the plot of Diffusion constant vs. the enzyme turnover rate is essentially the stroke size<sup>4</sup>. The stroke size for catalase determined from the measured values of  $\sqrt{\alpha}$  is as large as 31.6nm. This discrepancy between the hydrodynamic model and the experiments<sup>10</sup>, argues that a careful examination of other possible channels of fluorescence correlation decay besides enhanced enzyme diffusion must be considered.

We wish now to show explicitly that due to damping the motions of a swimming enzyme cannot excite a significant acoustic response in the fluids as was envisioned in the reaction heat mechanism proposed by Riedel et al. To treat the hydrodynamic effects beyond the low Mach number regime, one only needs to replace the long-time limit of the friction  $\zeta_0$  (given in our previous analysis using only steady hydrodynamics) with its frequency-dependent counterpart  $\zeta(\omega)$ . The frequency-dependent friction coefficient accounts both for finite momentum diffusivity and for sound propagation. Wolynes and McCammon<sup>13</sup> calculated  $\zeta(\omega)$  for a biopolymer decades ago by using nonsteady hydrodynamics. To focus on the acoustic effects alone one simply needs to omit the viscous contribution from their expression for  $\zeta(\omega)$ . This simplification gives a pure acoustic drag coefficient at zero-frequency with the value  $\zeta(0) = \zeta_0 \frac{a c_0}{9 \nu_0}$ . When substituted again into the long-time limit equation of motion to obtain the swimming velocity we find the pure acoustic mode would be even more strongly coupled to the motion of the swimmer than the viscous modes are. Thus if sound plays a significant role, the enzyme motion would be actually more strongly damped than it was our calculation where only the viscous coupling is taken into account. The acoustic effects clearly reduce the swimming speed: the swimming velocity when the pure acoustic effect is dominant turns out to be  $v_{sound} = \frac{9 \nu_0}{a c_0} v_{swim}$ . That the acoustic effect is small and that it does not influence the locomotion is not unexpected. The perturbed pressure field usually relaxes



more rapidly to its steady value than the perturbed velocity field such that the prefactor of  $v_{sound}$  as  $\frac{\nu_0}{ac_0} = \frac{a/c_0}{a^2/\nu_0} = \frac{\tau_s}{\tau_\nu}$  is also small, where  $\tau_s$  and  $\tau_\nu$  are the typical time for sound wave to propagate a distance of  $a$  and momentum to diffuse over an area of  $a^2$  respectively. That  $\tau_s$  is much smaller than  $\tau_\nu$  supports our prediction that compression indeed should be a small correction to the locomotion speed.

### III. CONCLUSION

While we see that individual enzymes can swim randomly when they carry out a catalytic reaction and that therefore chemistry can in principle lead to enhanced diffusion of an individual enzyme proportional to the enzyme's activity, the effect should be quite a bit smaller than the recent experiments seem to suggest. We therefore think alternate explanations of the data need to be entertained. The absolute measured changes of the fluorescence correlations are quite subtle, so other ways of losing fluorescence correlation other than through enhanced diffusion out of the illuminated region should be considered. One possible way to lose fluorescence correlation is for an intermediate in the catalysis or a reaction product of the enzymatic reaction to quench the fluorescence when an active species forms transiently near the monitored fluorophore. In that case an extra source of decorrelation would be present and it would be proportional to the enzyme activity just as the proposed diffusion enhanced effect is. It seems essential to rule out this possibility for the alkaline phosphatase system which by catalyzing the hydrolysis of nitrophenylphosphate yields a fluorescently active product nitrophenol. Likewise the two enzymes urease and catalase have spectroscopically and electronically active metal centers. As these centers undergo the catalytic cycle they might yield species that can transiently quench the fluorescence, most likely through electron transfer. In order to estimate the size of these quenching effects much detailed spectroscopic electrochemical and kinetic data would be needed. This is beyond the scope of our present effort. It is most interesting, however, that the one system that displayed no excess decorrelation, triose phosphate isomerase, also lacks metal centers and catalyzes a reaction that would involve no active products. For this system then the activity induced quench mechanism could not operate and thus would explain why no enhanced diffusion was observed in that case. We point out that the excess transient quenching

mechanism can be experimentally distinguished from the enhanced diffusion mechanism by carefully measuring the effect of changing the size of the illuminated region. Such a change should alter the diffusion signal but not the transient quenching signal.

As we finished preparing this account for publication, a theoretical study by Golestanian<sup>28</sup> appeared that supports the idea that stochastic swimming can contribute to the enzyme activity enhancement of diffusion, but like us he concludes that the effects should be much smaller than those that were measured. In seeking an explanation of the observation he points out that the global heating by the enzyme reaction could result in heating up the whole sample which would appear as enhanced diffusion. The original authors of the experiments had considered but ruled out such an effect in their work. Obviously besides checking out the possibility of excess transient quenching we hope that future experiments with accurate temperature monitoring will also be undertaken. We look forward to such experimental investigations.

## ACKNOWLEDGMENTS

We thank Jose Onuchic for a most helpful discussion. This work was supported by the Center for Biologically Physics sponsored by National Science Foundation Grants PHY-1308264 and PHY-1427654. Additional support was provided by the D. R. Bullard-Welch Chair at Rice University Grant C-0016.

## Appendix A: Path-integral Treatment for Brownian Crossing of a Inverted Harmonic Barrier

Brownian dynamics within the inverted harmonic potential  $U(x) = -\frac{1}{2}m\omega^{\ddagger 2}x^2 + \Delta U^{\ddagger}$  is described by an overdamped Langevin equation:

$$\frac{dx}{dt} = bx + \eta(t), \quad \langle \eta(t)\eta(t') \rangle = 2D\delta(t - t') \quad (\text{A1})$$

where  $b = \beta D m \omega^{\ddagger 2}$  is the frequency involved in overdamped motion. To describe the probability functional for a Brownian path one has the Onsager-Marchlup Lagrangian associated

with the Brownian motion in the inverted harmonic potential<sup>29</sup>:

$$\mathcal{L}(\dot{x}, x, t) = \frac{1}{4D}(\dot{x} - bx)^2 + \frac{b}{2} \quad (\text{A2})$$

The corresponding Euler-Lagrange equation for the dominant path  $\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = 0$  then is:

$$\ddot{x} - b^2x = 0 \quad (\text{A3})$$

This equation of motion for  $x$  describes an unstable damped motion. Given the initial conditions ( $t = 0$ ) and the final positions ( $t = T$ ) of the diffusing degree of freedom  $x$  as  $x(0)$  and  $x(T)$  respectively, one finds the most probable path for  $x(t)$ :

$$x(t) = \frac{x(T) \sinh(bt) + x(0) \sinh[b(T-t)]}{\sinh(bT)} \quad (\text{A4})$$

The most probable leaping path starts from the initial position  $x(0) = -\sqrt{\frac{2\Delta U^\ddagger}{m\omega^\ddagger{}^2}}$  proceeding to the final position  $x(T) = +\sqrt{\frac{2\Delta U^\ddagger}{m\omega^\ddagger{}^2}}$ . At these endpoints the free energy has fallen by  $\Delta U^\ddagger$  from the barrier top. Thus the dominant leaping path for traversing in a time  $T$  is given by:

$$x(t) = \sqrt{\frac{2\Delta U^\ddagger}{m\omega^\ddagger{}^2}} \frac{\sinh(bt) - \sinh[b(T-t)]}{\sinh(bT)} \quad (\text{A5})$$

## Appendix B: Transit time in 1D inverted harmonic potential

From the dominant path equation (A4), one can calculate the conditional transition probability density  $\Phi(x, t|x_0, 0)$ . This is also the propagator of a 1D Fokker-Planck equation with absorbing boundary conditions at the start and end of the trajectory at the bottom of the inverted well:

$$\Phi(x, t|x_0, 0) = \int \mathcal{D}[x(t)] \exp \left[ - \int_0^t d\tau \mathcal{L}(\dot{x}, x, \tau) \right] \quad (\text{B1})$$

$$= \left[ \frac{b}{2\pi D(e^{2bt} - 1)} \right]^{\frac{1}{2}} \exp \left[ - \frac{b(x - x_0 e^{bt})^2}{2D(e^{2bt} - 1)} \right] \quad (\text{B2})$$

Recalling that the Fokker-Planck equation is a probability balance equation we see the probability flux for starting at  $x_0$  at  $t = 0$  while ending up at  $x$  at time  $t$  is proportional to the probability of the transit time being  $t$ :

$$\frac{\partial \Phi(x, t|x_0, 0)}{\partial t} = - \frac{\partial}{\partial x} J(x, t|x_0, 0) \quad (\text{B3})$$

$$J(x, t|x_0, 0) = -D e^{\beta U(x)} \frac{\partial}{\partial x} [e^{-\beta U(x)} \Phi(x, t|x_0, 0)] \quad (\text{B4})$$

We will denote the initial and final position  $\pm\sqrt{\frac{2\Delta U^\ddagger}{m\omega^\ddagger{}^2}}$  as  $\pm\Delta x$  for convenience. Thus the reactive flux  $J(\Delta x, T | -\Delta x, 0)$  with transit time  $T$  is:

$$J(\Delta x, T | -\Delta x, 0) = \frac{b\Delta x}{4} \sqrt{\frac{b}{\pi D}} \frac{\exp[-\frac{b\Delta x^2}{2D} \coth(bt/2)]}{\sinh(bt/2) \sqrt{\sinh(bt)}} \quad (\text{B5})$$

The probability distribution density  $P(t)$  for transit time is just the normalized reactive flux:

$$P(t) = \frac{J(\Delta x, t | -\Delta x, 0)}{\int_0^\infty J(\Delta x, t | -\Delta x, 0) dt} \quad (\text{B6})$$

$$= \frac{b\Delta x}{2[1 - \text{erf}(\Delta x \sqrt{b/(2D)})]} \sqrt{\frac{b}{\pi D}} \frac{\exp[-\frac{b\Delta x^2}{2D} \coth(bt/2)]}{\sinh(bt/2) \sqrt{\sinh(bt)}} \quad (\text{B7})$$

$$= \frac{b\sqrt{\beta\Delta U^\ddagger}}{1 - \text{erf}(\sqrt{\beta\Delta U^\ddagger})} \frac{\exp[-\frac{b\Delta x^2}{2D} \coth(bt/2)]}{\sinh(bt/2) \sqrt{\sinh(bt)}} \quad (\text{B8})$$

The mean transit time  $T$  can now be obtained as the first moment of  $P(t)$  in the long-time limit ( $bt \gg 1$ ):

$$\begin{aligned} T &= \int_0^\infty tP(t) dt \\ &\approx \int_0^\infty \tau d\tau \frac{b\sqrt{\beta\Delta U^\ddagger}}{e^{-\beta\Delta U^\ddagger}/\sqrt{\pi\beta\Delta U^\ddagger}} \times \frac{\exp[-\beta\Delta U^\ddagger(1 + 2e^{-bt})]}{e^{bt/2}/2 \times \sqrt{\pi e^{bt}}} \\ &= \int_0^\infty \tau d\tau \times 2b\beta\Delta U^\ddagger \exp[-b\tau - 2\beta\Delta U^\ddagger e^{-b\tau}] \\ &\approx \frac{1}{b} \ln[(2\beta\Delta U^\ddagger) + \gamma] \quad (\text{when } \beta\Delta U^\ddagger \gg 1) \end{aligned}$$

where two approximations have been made: i), the free-energy barrier is high such that  $\beta\Delta U^\ddagger \gg 1$ ; ii),  $\coth(x) = 1 + \frac{2e^{-x}}{1-e^{-x}} \approx 1 + 2e^{-x}$  when  $x$  is large. From the distribution function, one can also calculate that the most probable transit time  $t_{m.p.}$ , which differs from mean transit time  $T$  by the term containing the Euler constant:  $t_{m.p.} = \frac{1}{b} \ln(2\beta\Delta U^\ddagger)$ .

## Appendix C: The most probable traversal path for a composite system: the model swimming enzyme

We start with the equation of motion for the coupled Brownian dynamics for the three domains of the enzyme:

$$\frac{dr_i}{dt} = \beta \sum_j D_{ij} F_j + \eta_i(t) \quad (C1)$$

$F_i$  is the effective spring force exerted on each sphere and is determined by the excited potential  $F_e$ .  $\eta_i(t)$  are white noise terms that are related to the multiparticle diffusion tensor  $D_{ij}$  by the fluctuation-dissipation theorem  $\langle \eta_i(t) \eta_j(t') \rangle = 2D_{ij} \delta(t - t')$ . At the Oseen level the diffusion tensor  $D_{ij}$  is  $D_{ij} = k_B T [\delta_{ij} \zeta_0^{-1} + (1 - \delta_{ij}) T(r_{ij})]$ , where  $T(r) = (8\pi\eta_0 r)^{-1} (I + \frac{\hat{r}\hat{r}}{r^2})$  is the Oseen Tensor. This set of equations can be expressed more explicitly as:

$$\frac{d}{dt} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \frac{1}{\zeta_0} \begin{pmatrix} 1 & \frac{3a}{2x} & \frac{3a}{4x} \\ \frac{3a}{2x} & 1 & \frac{3a}{2x} \\ \frac{3a}{4x} & \frac{3a}{2x} & 1 \end{pmatrix} \left[ \begin{pmatrix} k & -k & 0 \\ -k & 0 & k \\ 0 & k & -k \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} + \begin{pmatrix} kx \\ -2kx \\ kx \end{pmatrix} \right] + \begin{pmatrix} \eta_1(t) \\ \eta_2(t) \\ \eta_3(t) \end{pmatrix} \quad (C2)$$

$$= \frac{k}{\zeta_0} \begin{pmatrix} 1 - \frac{3a}{2x} & -1 + \frac{3a}{4x} & \frac{3a}{4x} \\ -1 + \frac{3a}{2x} & 0 & 1 - \frac{3a}{2x} \\ -\frac{3a}{4x} & 1 - \frac{3a}{4x} & -1 + \frac{3a}{2x} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} + \frac{k}{\zeta_0} \begin{pmatrix} x - \frac{9}{4}a \\ -2x + 3a \\ x - \frac{9}{4}a \end{pmatrix} + \begin{pmatrix} \eta_1(t) \\ \eta_2(t) \\ \eta_3(t) \end{pmatrix} \quad (C3)$$

$$= \mathbf{A} \mathbf{r} + \mathbf{b} + \boldsymbol{\eta}(t) \quad (C4)$$

Diagonalizing the composite force-mobility matrix part leads us to three real eigenvalues associated with three eigen-modes  $r'_1$ ,  $r'_2$  and  $r'_3$ . These eigenvalues are also the three eigenfrequencies of the swimmer's dynamics. By replacing the spring constraint  $b$  in the 1D traversal time path with the eigenvalues, the dominant path for these independent eigenmodes  $r'_1$ ,  $r'_2$  and  $r'_3$  can be found as well as the paths for the original coordinates  $r_1$ ,  $r_2$  and  $r_3$ . The resulting eigenvalues and eigenvectors are summarized as follows:

**Eigenvalues:**  $\lambda_1 = 0$ ,  $\lambda_2 = -\sqrt{3} \frac{k}{\zeta_0} (1 - \frac{5a}{4x})$ ,  $\lambda_3 = \sqrt{3} \frac{k}{\zeta_0} (1 - \frac{5a}{4x})$

**Eigenvectors:**  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ,  $\begin{pmatrix} -2 + \sqrt{3} + \frac{a}{x}(3 - 2\sqrt{3}) \\ 1 - \sqrt{3} + \frac{a}{x}(\frac{-9+5\sqrt{3}}{4}) \\ 1 \end{pmatrix}$ ,  $\begin{pmatrix} -2 - \sqrt{3} + \frac{a}{x}(3 + 2\sqrt{3}) \\ 1 + \sqrt{3} + \frac{a}{x}(\frac{-9-5\sqrt{3}}{4}) \\ 1 \end{pmatrix}$

The eigenmodes  $r'_i(t)$  can be transformed to the original coordinates  $r_i(t)$  through a transformation matrix  $\mathbf{P}$  composed of the eigenvectors:

$$\begin{pmatrix} r'_1 \\ r'_2 \\ r'_3 \end{pmatrix} = \mathbf{P}^{-1} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 + \sqrt{3} + \frac{a}{x}(3 - 2\sqrt{3}) & -2 - \sqrt{3} + \frac{a}{x}(3 + 2\sqrt{3}) \\ 1 & 1 - \sqrt{3} + \frac{a}{x}(\frac{-9+5\sqrt{3}}{4}) & 1 + \sqrt{3} + \frac{a}{x}(\frac{-9-5\sqrt{3}}{4}) \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \quad (\text{C5})$$

$$= \begin{pmatrix} \frac{1}{3} + \frac{a}{12x} & \frac{1}{3} - \frac{a}{6x} & \frac{1}{3} + \frac{a}{12x} \\ -\frac{1}{6} - \frac{(1-3\sqrt{3})a}{24x} & -\frac{1+\sqrt{3}}{6} + \frac{(1-\sqrt{3})a}{12x} & \frac{2+\sqrt{3}}{6} - \frac{(1+\sqrt{3})a}{24x} \\ -\frac{1}{6} - \frac{(1+3\sqrt{3})a}{24x} & -\frac{1-\sqrt{3}}{6} + \frac{(1+\sqrt{3})a}{12x} & \frac{2-\sqrt{3}}{6} - \frac{(1-\sqrt{3})a}{24x} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \quad (\text{C6})$$

The collective mode  $r'_1$  with zero eigenvalue is apparently the motion for center of mass at the Oseen limit ( $\frac{a}{x} \rightarrow 0$ ):  $r'_1 = \frac{1}{3}(r_1 + r_2 + r_3)$ . Transformation of the constant part of equation (C3) gives:

$$\mathbf{b}' = \mathbf{P}^{-1}\mathbf{b} = \begin{pmatrix} -\frac{7ka^2}{8\zeta_0 x} \\ \frac{kx}{6\zeta_0}(3 + 3\sqrt{3}) - \frac{ka}{8\zeta_0}(9 + 5\sqrt{3}) + \frac{7ka^2}{16\zeta_0 x}(1 - \sqrt{3}) \\ \frac{kx}{6\zeta_0}(3 - 3\sqrt{3}) - \frac{ka}{8\zeta_0}(9 - 5\sqrt{3}) + \frac{7ka^2}{16\zeta_0 x}(1 + \sqrt{3}) \end{pmatrix} \quad (\text{C7})$$

The swimming velocity is the first entry of transformed constant vector:  $v_{swim} = b'_1 = -\frac{7ka^2}{8\zeta_0 x}$ . Obviously if we set the initial position of  $r'_1$  as  $x$ , its final position will be the stroke size  $x + L$  at transit time  $T$ . Thus the dominant path for the center of mass is:

$$r'_1(t) = -\frac{7ka^2}{8\zeta_0}t + x = -\frac{L}{T}t + x \quad (\text{C8})$$

Equation (C8) also reveals that we can estimate  $L$  once we have obtained transit time  $T$ . The dominant paths for  $r'_2(t)$  and  $r'_3(t)$  follow equation (A4) by substituting in the corresponding eigenvalue and boundary conditions.

## Appendix D: Transit time for the model swimmer

Since we are especially interested in the motion along the unstable mode  $r'_3(t)$  that gives rise to the transit time of the catalytic reaction, we must find the boundary conditions

$r'_3(0)$  and  $r'_3(T)$ . The exponential escape motion at the transition state is given by a rate  $\lambda_3$ . Motion along the eigen coordinate is not parallel to the displacement that determines the free energy barrier  $\Delta U^\ddagger$ . We thus need to calculate the angle  $\phi$  between the energetic reaction coordinate and the unstable mode in the euclidean space spanned by  $r_1$ ,  $r_2$  and  $r_3$ . The direction of unstable mode is proportional to the components of the eigenvector  $\mathbf{e}_3$  associated with eigenvalue  $\lambda_3$  along the reaction coordinate is  $\mathbf{e}_{rxn} = \frac{1}{\sqrt{2}}(-1, 1, 0)$ . We find the angle is:

$$\begin{aligned}\cos \phi &= \frac{1}{\sqrt{2}} \frac{(-1, 1, 0) \cdot (-2 - \sqrt{3} + \frac{a}{x}(3 + 2\sqrt{3}), 1 + \sqrt{3} + \frac{a}{x}(\frac{-9-5\sqrt{3}}{4}), 1)}{|(-2 - \sqrt{3} + \frac{a}{x}(3 + 2\sqrt{3}), 1 + \sqrt{3} + \frac{a}{x}(\frac{-9-5\sqrt{3}}{4}), 1)|} \\ &\approx \frac{1}{4\sqrt{2}(3 + \sqrt{3})} [12 + 8\sqrt{3} - (21 + 13\sqrt{3})\frac{a}{x}]\end{aligned}$$

Thus the transit time for the swimmer with proper traversal barrier is:

$$T = \frac{\ln(2\beta V^\ddagger) + \gamma}{\lambda_3} = \frac{\ln(2\beta \Delta U^\ddagger) + \gamma + \ln[\sqrt{3}(1 - \frac{5a}{4x})/\cos^2 \phi]}{\lambda_3} \quad (\text{D1})$$

## Appendix E: Calculation of stroke size

The product of transit time  $T$  expressed in equation (D1) of swimming speed immediately leads us to the approximation for the stroke size  $L$  quoted in the main text:

$$L = \frac{\ln(2\beta \Delta U^\ddagger) + \gamma + \ln[\sqrt{3}(1 - \frac{5a}{4x})/\cos^2 \phi]}{\lambda_3} \times \frac{7ka^2}{8\zeta_0 x} \quad (\text{E1})$$

$$= \frac{7a^2[\ln(2\beta \Delta U^\ddagger) + \gamma + \ln(\sqrt{3}/\cos^2 \phi)]}{8\sqrt{3}x} \quad (\text{E2})$$

## REFERENCES

- <sup>1</sup>J. Howard, “Mechanics of Motor Proteins and the Cytoskeleton” Sinauer Press, Sunderland MA, (2001).
- <sup>2</sup>C. Veigel and C.F. Schmidt, Nat. Rev. Mol. Cell Bio. **12**, 163 (2011).
- <sup>3</sup>T. Shen and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **101**, 8547 (2004).
- <sup>4</sup>S. Wang and P. G. Wolynes, J. Chem. Phys. **135**, 051101 (2011).

- <sup>5</sup>S. Wang and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **108**, 15184 (2011).
- <sup>6</sup>S. Wang and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **109**, 6446 (2012).
- <sup>7</sup>S. Wang and P. G. Wolynes, J. Chem. Phys. **136**, 145102 (2012).
- <sup>8</sup>S. Wang and P. G. Wolynes, J. Chem. Phys. **138**, 12A521 (2013).
- <sup>9</sup>S. Wang and P. G. Wolynes, J. Chem. Phys. **139**, 235103 (2013).
- <sup>10</sup>C. Riedel, R. Gabizon, C. A. M. Wilson, K. Hamadani, K. Tsekouras, S. Marqusee, S. Presse and C. Bustamante Nature **517**, 227 (2015).
- <sup>11</sup>H. S. Muddana, S. Sengupta, T. E. Mallouk, A. Sen and P. J. Butler J. Am. Chem. Soc. **132**, 2110 (2010).
- <sup>12</sup>J. A. McCammon, B. R. Gelin, M. Karplus and P. G. Wolynes Nature **262**, 325 (1976).
- <sup>13</sup>J. A. McCammon and P. G. Wolynes, J. Phys. Chem. **66**, 452 (1977).
- <sup>14</sup>E.M. Purcell, Am. J. Phys. **45**, 3 (1977).
- <sup>15</sup>A. Shapere and F. Wilczek, Phys. Rev. Lett. **58**, 2051 (1987).
- <sup>16</sup>E. Lauga, Soft Matt. **7**, 3060 (2011).
- <sup>17</sup>J.M. Yeomans, D.O. Pushkin and H. Shum, Eur. Phys. J. **223**, 1771 (2014).
- <sup>18</sup>A. Najafi and R. Golestanian Phys. Rev. E **69**, 062901 (2004).
- <sup>19</sup>C. Hyeon and J. N. Onuchic Proc. Natl. Acad. Sci. U.S.A. **104**, 17382 (2007).
- <sup>20</sup>W. Paxton, S. Sundararajan, T. Mallouk and A. Sen, A. Angew. Chem. ,Int. Ed. **45**, 5420 (2006).
- <sup>21</sup>K. Itoh and M. Sasai Proc. Natl. Acad. Sci. U.S.A. **107**, 7775 (2010).
- <sup>22</sup>P. G. Wolynes and J. M. Deutch J. Phys. Chem. **67**, 733 (1977).
- <sup>23</sup>G. Hummer, J. Phys. Chem. **120**, 516 (2004).
- <sup>24</sup>B. W. Zhang, D. Jasnow and D. M. Zuckerman J. Phys. Chem. **126**, 074504 (2007).
- <sup>25</sup>S. Chaudhury and D. E. Makarov, J. Phys. Chem. **133**, 034118 (2010).
- <sup>26</sup>H. S. Chung, J. M. Louis and W. A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **106**, 11873 (2009).
- <sup>27</sup>H. S. Chung, K. McHale, J. M. Louis, W. A. Eaton Science **335**, 981 (2012).
- <sup>28</sup>R. Golestanian arXiv:1508.03219 [<http://arxiv.org/abs/1508.03219>].
- <sup>29</sup>H. S. Wio “Path Integrals for Stochastic Processes - An introduction” World Scientific Publishing Co. Pte. Ltd., Singapore, (2003).



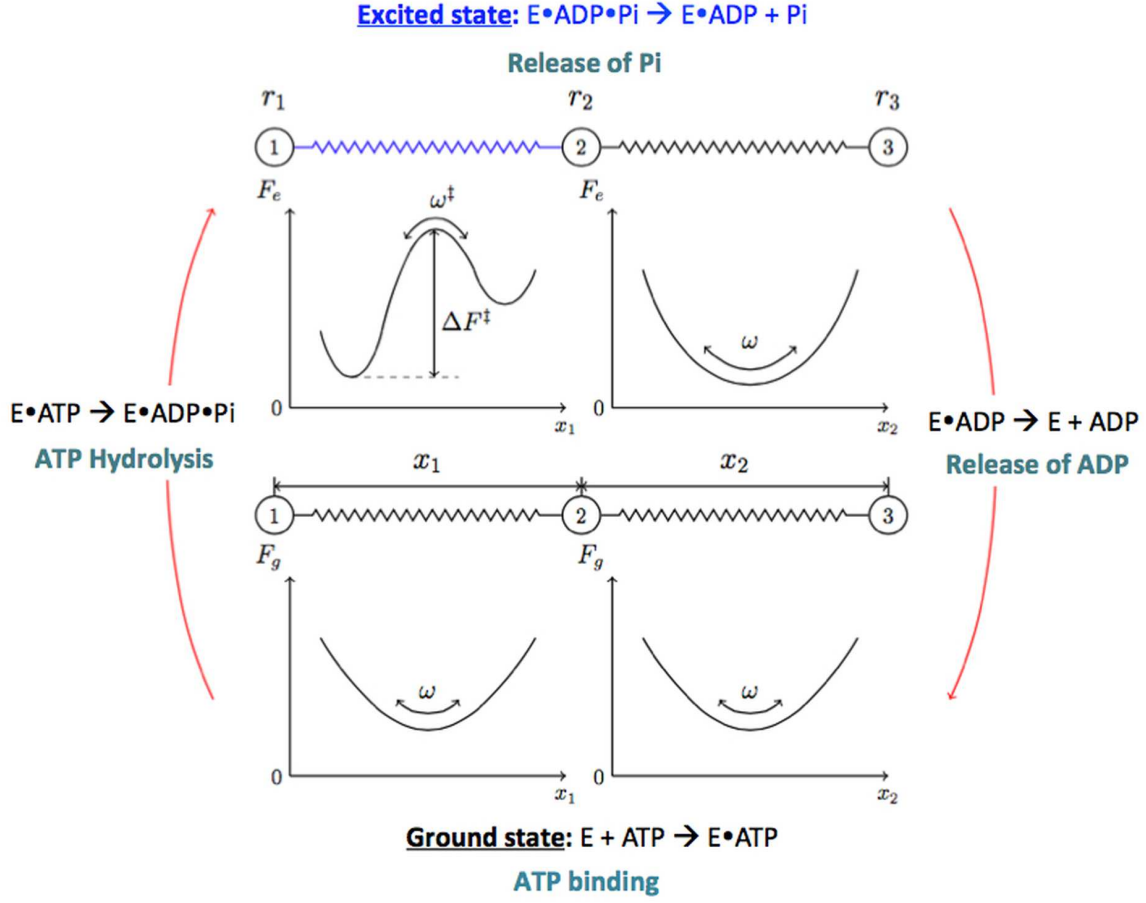


FIG. 1. Illustration of a 3-sphere model swimmer.  $r_1$ ,  $r_2$  and  $r_3$  are coordinates for the three spheres in real 3D space. The panel above shows the free energy profile of excited state  $F_e$  that projected onto internal coordinates  $x_1$ (left) and  $x_2$ (right), while the panel below shows the profile for the ground state  $F_g$ . All the curvatures are taken to be the same,  $\omega = \sqrt{k/m}$ . The instantaneous transition processes are denoted by the two red arrows. All the relevant chemical steps are also indicated.